

**Comparison of proteomic and genomic analyses of the human breast cancer cell line T47D and the antiestrogen-resistant derivative T47D-r**

Martina Huber<sup>1</sup>

Inke Bahr<sup>1</sup>

Jörn R. Krätzschar

Andreas Becker

Eva-Christina Müller<sup>2</sup>

Peter Donner

Hans-Dieter Pohlenz

Martin R. Schneider

Anette Sommer<sup>3</sup>

Research Laboratories of Schering AG, Berlin, Germany

<sup>1</sup>The first two authors contributed equally

<sup>2</sup>Protein Chemistry Group, Max-Delbrück-Center for Molecular Medicine, 13092 Berlin, Germany

<sup>3</sup>Corresponding author:           Anette Sommer  
  Schering AG  
  Enabling Technologies  
  13342 Berlin  
  Germany  
  Tel.: +49-30-468-17974  
  Fax: +49-30-468-16707  
  E-mail: anette.sommer@schering.de

**Running title:** Comparison of proteomic and genomic analysis of T47D and -r

## Abbreviations Page:

2D, two-dimensional; BSA, bovine serum albumin; CCS; charcoal-treated fetal calf serum; CHAPS, 3-((3-cholamidopropyl)dimethylammonio)-1-propane sulfonate; cRNA, complementary RNA; DMSO, dimethyl sulfoxide; DTT, dithiothreitol; E<sub>2</sub>, 17β-estradiol; ECL, enhanced chemiluminescence; ER, estrogen receptor; ESI, electrospray ionization; EST, expressed sequence tag; FCS, fetal calf serum; IEF, isoelectric focussing; IPG, immobilized pH gradient; MALDI, Matrix-assisted laser desorption/ionization; MS, mass spectrometry; MTT, 4,5-[dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide; PAGE, polyacrylamide gel electrophoresis; PBS, phosphate-buffered saline; PCR, polymerase chain reaction; PMSF, phenylmethylsulfonyl fluoride; PSD, post source decay; SSC, sodium chloride / sodium citrate; TFA, trifluoroacetic acid; TOF, time-of-flight; ZM, ZM 182780.

## Summary

In search of novel mechanisms leading to the development of antiestrogen-resistance in human breast tumors, we analyzed differences in the gene and protein expression pattern of the human breast carcinoma cell line T47D and its derivative T47D-r, which is resistant toward the pure antiestrogen ZM 182780 (Faslodex<sup>TM</sup>, fulvestrant. Affymetrix DNA chip hybridizations on the commercially available HuGeneFL and Hu95A arrays were carried out in parallel to the proteomics analysis where the total cellular protein content of T47D or T47D-r was separated on 2D-gels. 38 proteins were found to be reproducibly up- or down-regulated more than 2-fold in T47D-r versus T47D in the proteomics analysis. Comparison with differential mRNA analysis revealed that 19 of these were up- or down-regulated in parallel with the corresponding mRNA molecules, among those are the protease cathepsin D, the GTPases Rab11a and MxA, and the secreted protein hAG-2. For nine proteins the corresponding mRNA was not found to be differentially expressed and for ten proteins an inverse regulation was found at the mRNA level. In summary, mRNA expression data, when combined with proteomic information, provide a more detailed picture of how breast cancer cells are altered in their antiestrogen-resistant compared to the antiestrogen-sensitive state.

## Introduction

Endocrine therapy plays an important role in the management of breast cancer at various stages (1). In the clinic, tamoxifen is the most widely used non-steroidal antiestrogen (2). It is effective in first-line and adjuvant treatments. However, the fact that tamoxifen has partial agonistic activity in the uterus and that long-term treatment was shown to result in tamoxifen-resistant breast tumors, led to the development of the pure antiestrogen ZM 182780 (fulvestrant, Faslodex<sup>TM</sup>, ICI 182780) which lacks ER alpha agonistic activity in all tissues, decreases the half-life of the ER alpha and is still effective after tamoxifen failure (3). Fulvestrant has been tested in phase III clinical trials versus tamoxifen for first line therapy of advanced breast cancer and versus the aromatase inhibitor anastrozole in patients with tamoxifen-refractory breast cancer (4). Fulvestrant was at least as effective and equally well tolerated as anastrozole for the treatment of postmenopausal women with advanced and metastatic breast cancer and gained approval from the US Food and Drug Administration in 2002 (2). In cell culture and xenograft models resistance toward ZM 182780 occurs later than after tamoxifen treatment; therefore ZM 182780 may provide a longer duration of response compared to tamoxifen (5, 6).

Breast tumors as well as breast cancer cell lines have the capacity to adapt to the prevailing concentration of antiestrogen, resulting in resistance to the drug. The mechanisms which lead to antiestrogen-resistance in breast cancer patients as well as in cell culture model systems are still poorly understood to date (7).

Therefore, we established an antiestrogen-resistant breast cancer cell line as an *in vitro* model system in order to analyze differentially expressed genes and proteins. To this end, the human ER alpha-positive breast carcinoma cell line T47D, which is growth-stimulated by E<sub>2</sub> and sensitive to antiestrogens, was exposed to the pure antiestrogen ZM 182780 (Faslodex<sup>TM</sup>) at a concentration of 10<sup>-6</sup> M over a period of several months. From this treatment a ZM 182780-resistant cell line evolved which was called T47D-r (8). We demonstrated that T47D-r cells have lost expression of the ER alpha and show an up-

regulation of the EGF receptor and c-erb B2, supporting the notion that cross-talk exists between ER alpha and growth factor signaling pathways (8).

The screening of global mRNA changes by hybridization of RNA to either DNA or oligonucleotide arrays in a high throughput manner has only become possible in recent years (9) and has been applied to the analysis of gene expression profiles of different breast cancer cell lines, to the study of the response of breast cancer cells to E<sub>2</sub> and to the molecular profiling of breast cancer biopsies in relation to the estrogen receptor alpha status and to the successful subclassification of human breast tumors (10-15).

Proteomics is a complementary tool to mRNA analysis for assessing global changes in cellular protein expression and it can provide additional insight into post-translational modifications of proteins. Recent progress in the field of proteomics has resulted in publications on proteomes for normal and tumor-derived breast cancer cell lines and breast cancer biopsies and results have been reviewed and discussed (16-21). 2D-electrophoresis coupled with MS identification of proteins has for example been carried out for establishing a catalogue of proteins which are expressed in human tumorigenic and non-tumorigenic breast cell lines (22-27). Although cultivated cells tend to have a lower complexity of expressed proteins than human biopsies, comparison of only distantly related or rather unrelated cellular systems resulted in a large number of differentially expressed proteins – however with limited clinical relevance. The proteomics studies of normal human breast, pre-malignant lesions, and breast cancer biopsies generated data with stronger clinical impact but even higher biological complexity (28-32). Moreover, in specialized approaches luminal and myoepithelial breast cells from reduction mammoplasties were analyzed by proteomics after a double antibody magnetic sorting technique (33) and membrane proteins of breast cancer cell lines were studied after cell fractionation (34). This lowered the complexity of the systems and increased the number of low-abundance proteins that could be detected.

Recently, also a proteomics study of the human breast cancer cell line MCF-7 with or without treatment of the anti-tumor drug doxorubicin has been reported (35).

Today, only a limited number of studies have been published in which both technologies - proteomics and mRNA gene expression profiling - have been compared using the same samples (36-38). There are reports that absolute protein and mRNA expression levels do not correlate well when total expression profiles are analyzed (39). With regard to a comparison of those genes and proteins which are differentially expressed several studies have been published recently (40-47). The conclusion drawn by the authors on the correlation of differential RNA and protein expression varies between moderate and no correlation and seems at least partially to depend on the abundance of the mRNA or the protein under investigation.

The aim of the present study was to define and compare differential gene and protein expression in a model system of antiestrogen-sensitive and -resistant breast cancer and to evaluate concordance of mRNA and protein data. It is expected that a more detailed analysis of these changes will give new hints about the mechanisms which lead to the development of antiestrogen-resistant breast cancer and may eventually help to propose novel strategies and targets for an effective breast cancer therapy.

## **Experimental Procedures**

**Chemicals** – The steroid hormone 17 $\beta$ -estradiol (E<sub>2</sub>) and ZM 182780, synthesized in the Laboratories of Schering AG (Berlin, Germany), were solubilized in absolute ethanol at a stock concentration of 10<sup>-3</sup> and 10<sup>-2</sup> M, respectively, and stored at -20°C. Dilutions were performed in absolute ethanol.

**Cell culture** – The human breast cancer cell line T47D was obtained from the American Type Cell Culture collection (Rockville, MD). The cells were routinely grown in phenol red-free RPMI 1640 medium (Biochrom, Berlin, Germany) supplemented with 10% FCS (LifeTechnologies, Karlsruhe, Germany), 2 mM glutamine (LifeTechnologies), 200 mU/ml insulin (Sigma, Deisenhofen, Germany) and  $10^{-10}$  M  $E_2$ . Prior to experiments, the cells were exposed to standard medium: phenol red-free RPMI 1640 medium, supplemented with 10% CCS, 2 mM glutamine and  $10^{-10}$  M  $E_2$ . The antiestrogen-resistant cells were grown continuously in standard medium with  $10^{-6}$  M ZM 182780. Cells were grown at 37°C in a humidified atmosphere with 5%  $CO_2$ . For generation of protein extracts,  $8 \times 10^6$  T47D were seeded in standard medium and T47D-r in standard medium plus  $10^{-6}$  M ZM 182780 (day 0). On day 3, medium was replaced for 24 hours with standard medium in case of T47D and standard medium plus  $10^{-7}$  M ZM 182780 in case of T47D-r. Cell cultures which had reached 70% confluence were harvested on day 4. For isolation of RNA, T47D were cultivated in standard medium and T47D-r in standard medium plus  $10^{-7}$  M ZM 182780 for 24 hours. polyA<sup>+</sup> RNA from two independent cell culture experiments was hybridized to the HuGeneFL array and total RNA from a further independent cell culture experiment was hybridized to the Hu95A array.

**Cell lines and culture conditions** – The human breast cancer cell lines T47D, ZR-75-1, and MCF-7 [48-50] were obtained from the American Type Cell Culture collection (Rockville, MD, USA). They were routinely grown in phenol red-free RPMI 1640 medium (Biochrom, Berlin, Germany) supplemented with 10% FCS from LifeTechnologies (Karlsruhe, Germany), 2 mM glutamine (Biochrom), 200 mU/ml insulin (Sigma, Taufkirchen, Germany) and  $10^{-10}$  M  $E_2$ . The parental cells were cultivated in parallel to the antiestrogen-resistant cells. Prior to an experiment, the cells were exposed to phenol red-free RPMI 1640 medium supplemented with 10% CCS, 2 mM glutamine and  $10^{-10}$  M  $E_2$ . The antiestrogen-resistant cells were grown continuously in phenol red-free RPMI 1640 medium supplemented with 10% CCS, 2 mM glutamine,  $10^{-10}$

M E<sub>2</sub>, and 10<sup>-6</sup> M ZM 182780. Cells were grown at 37°C in a humidified atmosphere with 5% CO<sub>2</sub>.

**Proliferation assay** – The proliferation rates of ZM 182780-sensitive T47D and ZM 182780-resistant T47D-r were measured with the MTT-assay. Cells grown to 70–80% confluence were harvested and seeded in 200 µl in 96 well microtiter plates at 2500 cells/well in phenol red-free RPMI 1640, 2 mM glutamine, 10% CCS (day zero). E<sub>2</sub> (10<sup>-10</sup> M), or the vehicle control ethanol at 0.001%, or E<sub>2</sub> (10<sup>-10</sup> M) plus ZM 182780 at concentrations of 10<sup>-10</sup> to 10<sup>-6</sup> M were added. Mitochondrial activity was measured on day seven after plating by adding 20 µl MTT reagent (Sigma, 5 mg/ml in PBS) for 3–5 hours at 37°C to the medium. Afterwards, the supernatant was removed and the formazan crystals were dissolved in DMSO. Absorbance was measured at 550 nm. Experiments were performed at least twice and one representative experiment with mean from six wells is shown.

**Generation of protein extracts** – For sample preparation 1 x 10<sup>7</sup>–1 x 10<sup>8</sup> cells were washed three times with PBS without Ca<sup>2+</sup>, Mg<sup>2+</sup> to remove serum proteins and detached by treatment with PBS (–Ca<sup>2+</sup>, –Mg<sup>2+</sup>) containing 2 mM EDTA. The pelleted cells were resuspended in PBS containing a protease inhibitor cocktail (Complete mini EDTA-free, pepstatin and phosphoramidon) (all from Roche, Mannheim, Germany) and Na-orthovanadate (Sigma) and stored in aliquots as dry pellets at –80°C before use. Cell pellets were resuspended in 20 mM Tris/HCl pH 8.0, 2 mM MgCl<sub>2</sub> containing protease inhibitor cocktail and Na-orthovanadate and lysed at 4°C with ultrasound (Sonorex RK31, Bandelin, Berlin, Germany) in the presence of glass beads of 1.7–2.0 mm diameter (Roth, Karlsruhe, Germany). The DNA in the lysate was digested with benzonase (Merck, Darmstadt, Germany) at room temperature for 30 min.

Afterwards, lysis buffer was added containing 7 M urea (Amersham Biosciences, Uppsala, Sweden), 2 M thiourea (Merck), 1% (w/v) DTT (Sigma), 2% (w/v) CHAPS (Electrophoresis Reagent, Sigma) and 0.5–1% IPG-buffer (Amersham Biosciences),



corresponding to the pH gradient. The lysate was vortexed for 1 min and centrifuged. The protein concentration was determined in the supernatant with the modified Bio-Rad protein assay kit (Bio-Rad, Hemel Hempstead, UK) and was 3–5 µg/µl.

**2D-SDS-PAGE** – Samples were loaded onto IPG strips (Immobiline DryStrips, Amersham Biosciences) using 100 µg total protein for analytical gels, which were silver-stained, and 1500 µg total protein for preparative gels, which were stained with colloidal Coomassie brilliant blue G250 (Serva, Heidelberg, Germany). For the analytical gels it was taken care that the amount of protein loaded onto the IPG strips was within the linear range of the silver stain detection. Spots which were saturated have not been included in the analysis. Isoelectric focussing was performed on an IPGphor unit (Amersham Biosciences) for IPG gradients pH 3–10, pH 5–8 (Bio-Rad) and pH 4–7 and on a Multiphor unit (Amersham Biosciences) for IPG gradients pH 6–11.

For the second-dimension SDS-PAGE using the Hoefer ISO-DALT system (Amersham Biosciences) focussed immobiline DryStrips were equilibrated in SDS buffer and transferred onto the surface of 13% T/2.6% C polyacrylamide/bisacrylamide gels (AppliChem, Darmstadt, Germany) and embedded in 1% agarose (Amersham Biosciences) containing traces of Bromophenol Blue (Amersham Biosciences). Gels were run in SDS-PAGE running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS) at 20°C until the dye-front reached the end of the gel. All samples were run at least in duplicates from three independent cell culture experiments leading to 12 gels for each pH gradient which were compared.

**Image analysis** – Analytical gels were scanned on an UMAX Mirage II scanner (UMAX systems, Düsseldorf, Germany) at 300 dpi with 12-bit gray scale levels in tagged image file format (TIFF) and images were analyzed using the Phoretix-2D advanced software V 5.1 (Nonlinear Dynamics Limited, Newcastle upon Tyne, UK). All gels in the analyses were scanned with identical parameters. In order to perform differential analysis of the 2D gels with the Phoretix-2D advanced software, the individual spots of each gel were

detected by their boundaries and the spot volume corresponding to the protein abundance was calculated automatically. Spot detection was performed with automatic splitting and features resulting from non-protein sources (e.g. dust particles and scratches) were filtered out. Manual editing of the gels was necessary especially in the higher molecular weight region. The background was subtracted using the mode-of-non-spot function. With the image warping function user seeds were set on gels for more accurate image superimposition. Spot matching was performed using T47D gel images as reference. Total spot volume normalization was used for accurate comparison of the different images. Only protein spots which were reproducibly different in all three experiments by at least a factor of two were considered to be differentially regulated and were excised from preparative gels for identification.

**Protein identification** – Spots excised from Coomassie-stained preparative gels were destained by washing first with digestion buffer (25 mM ammonium bicarbonate) and then with 50% acetonitrile/digestion buffer. These steps were repeated three times. Afterwards the gel pieces were dehydrated with acetonitrile and dried under vacuum in a SpeedVac concentrator. Digestion was performed with sequencing grade trypsin (Promega, Madison, WI, USA). A volume of 1–2  $\mu$ l of a solution containing 20 ng/ $\mu$ l trypsin in digestion buffer was added to each of the gel pieces. After 30 min of proteolysis at 37°C 10  $\mu$ l digestion buffer was added and digestion was continued overnight. The supernatant was concentrated and desalted on ZipTips (Millipore, Bedford, MA, USA) according to the manufacturer's instructions.

For MALDI-TOF and MALDI-PSD analysis the peptides were eluted with 1  $\mu$ l 50% acetonitrile/0.1% TFA onto the MALDI target. 1  $\mu$ l of a saturated  $\alpha$ -cyano-4-hydroxycinnamic acid solution in 70% acetonitrile/0.1% TFA was added to each sample. Spectra were acquired on a Voyager-DE STR (Perseptive Biosystems, Framingham, MA, USA). For nano-ESI-MS/MS analysis the peptides were eluted from the ZipTips with 3  $\mu$ l 50% acetonitrile/1% formic acid. The spectra were recorded on a Q-TOF

(Micromass, Manchester, UK) equipped with a nanoflow Z-spray ion source. The peptide mass fingerprints and MS/MS spectra obtained were analyzed searching the NCBI non-redundant (NCBI nr) protein database, the translated, assembled ESTs in the LifeSeqGold templates database (Incyte, trLGtemplates), and the non-redundant EST division of the GenBank, EMBL, DDBJ databases with the ProFound (<http://prowl.rockefeller.edu/cgi-bin/ProFound>) or Mascot (<http://www.matrixscience.com>) software.

**Western Blot analysis** – The cells were lysed on ice in lysis buffer containing 10 mM Tris, pH 7.5, 50 mM NaCl, 5 mM EDTA, 30 mM Na-pyrophosphate, 50 mM NaF, 1 mM PMSF (Sigma), and the protease inhibitor cocktail Complete<sup>TM</sup>-Mini (Roche), Phosphatase Inhibitor Cocktail I and II (Sigma), 1% (w/v) Triton-X-100 (Sigma). The plates were left on ice for 10 min and the lysates were transferred to Eppendorf tubes. The lysate was vortexed 2 x 30 sec and cellular debris was pelleted at 14.000 rpm for 20 min at 4°C. The supernatant was frozen at –80°C. Protein concentration of the cell lysate was determined with the DC Protein Assay Kit II (Bio-Rad). Equal amounts of protein were separated on 10% Bis-Tris gels (NuPAGE<sup>®</sup>, Invitrogen, Karlsruhe, Germany) and blotted to Optitran BA-S 85 reinforced nitrocellulose membrane (Schleicher & Schuell, Dassel, Germany) by semi-dry blotting in cathode buffer (25 mM Tris-base, 40 mM 6-amino hexane acid, 20% methanol), anode I buffer (30 mM Tris-base, 20% methanol), and anode II buffer (300 mM Tris-base, 20% methanol) at 0.8 mA/cm<sup>2</sup> for 60 min. The membrane was incubated in blocking buffer, i.e. Roti<sup>®</sup>-Block (Carl Roth GmbH, Karlsruhe, Germany) for 90 min at RT. Incubation with the primary antibodies was carried out in blocking buffer at 4°C overnight. The mouse monoclonal  $\alpha$ -cathepsin D antibody Ab-1 (clone BC011) was purchased from Oncogene<sup>TM</sup> Research Products (Boston, MA, USA) and  $\alpha$ - $\beta$ -actin was a rabbit polyclonal serum from Sigma. As secondary antibodies  $\alpha$ -mouse and  $\alpha$ -rabbit horseradish peroxidase-coupled antibodies (Amersham Biosciences) were used. The  $\alpha$ -cathepsin D antibody was diluted 1:500, the  $\alpha$ - $\beta$ -actin serum 1:1000 in blocking buffer. Afterwards, membranes were washed 3 x 5

min in PBS-T (PBS with 0.1% Tween-20) and incubated with the respective secondary horseradish-peroxidase coupled antibody (diluted 1:2000) for 1 h at RT. Membranes were again washed 3 x 5 min in PBS-T. Detection was carried out with the ECL Plus system (Amersham Biosciences) according to manufacturer's instructions. Blots were quantitated with the ImageMaster TotalLab 2.0 software (Amersham Biosciences).

**Affymetrix DNA chip hybridization** – 1–2  $\mu\text{g}$  of poly A<sup>+</sup> RNA or 10  $\mu\text{g}$  of total RNA were converted into double-stranded cDNA using a modified oligo-dT primer including a 5' T7 RNA polymerase promoter sequence and the Superscript Choice system for cDNA synthesis (LifeTechnologies). *In vitro* transcription was performed with T7 RNA polymerase (T7 Megascript kit, Ambion, Austin, TX, USA) and 0.5–1  $\mu\text{g}$  of double-stranded cDNA template in the presence of a mixture of ATP, CTP, GTP, biotin-11-CTP, and biotin-16-UTP (ENZO Diagnostics, Farmingdale, NY, USA). 20  $\mu\text{g}$  of cRNA were fragmented randomly by incubating in 40 mM Tris-acetate pH 8.1, 100 mM K<sup>+</sup> acetate, and 30 mM Mg<sup>2+</sup> acetate at 94°C for 35 min. The Affymetrix HuGeneFL array and the Hu95A array were hybridized, washed, and stained according to standard protocols. The arrays were scanned with a confocal scanner from Affymetrix. The hybridization intensities on the different chips were analyzed by an in-house developed software.

Affymetrix arrays were normalized using the Maximum Order Preserving Set method. This method identifies a subset of genes which do not change their expression between two experiments. The intensity scatter plot for all genes from this restricted set is approximated by a nonlinear regression function which is then used for normalization. In a first step all arrays are sorted by their intensity (median of all individual oligonucleotide signal intensities). The experiment with the median intensity of all arrays is chosen as the reference experiment. Each of the remaining arrays is normalized separately against the reference experiment. For each such pair-wise normalization all genes which are present on both arrays are identified. A gene is considered present if the p value for a one-sided matched-pairs-signed-rank Wilcoxon test applied to all unmasked perfect match (PM) versus mismatch (MM) oligonucleotide signal intensities is below 0.04. To identify a set

of oligo probe pairs belonging to genes with unchanged expression, all probe pairs for genes present on both arrays are sorted by their intensity differences of perfect match minus mismatch oligonucleotide (PM–MM) and are assigned the corresponding ranks. That is, each probe pair has a rank on the reference array and a usually different rank on the array to be normalized. Now, an iterative process is used to remove probe pairs with high rank jumps between the two experiments until a subset of probe pairs with identical ranks on both arrays is obtained (maximum order preserving set). The assumption is that probe pairs with high rank jumps belong to genes which do change their expression whereas probe sets with small rank jumps belong to genes with constant expression. For all probe pairs from the order preserving set the intensity difference PM–MM on the reference array is plotted against PM–MM on the array to be normalized and the best-fitting fourth order polynomial with no zero-order term is identified. In a last step all PM–MM differences from the array to be normalized are transformed by the thus identified nonlinear function. In case of the HuGeneFL array, where two independent hybridization experiments have been performed the average expression change between the replica experiments had to be at least 2-fold to be considered as regulated.

## Results

**Characterization of T47D and T47D-r** – T47D-r were generated from the antiestrogen-sensitive human breast cancer cell line T47D after continuous cultivation in presence of  $10^{-10}$  M  $E_2$  and  $10^{-6}$  M antiestrogen ZM 182780 for 31 weeks (8). Phase contrast photographs of T47D and T47D-r cells cultivated for 3 days in standard growth medium are shown (Fig. 1). Parental T47D and the ZM 182780-resistant derivative T47D-r have a similar morphology that is typical of tumor cells with an apolarized polygonal morphology.

The proliferative response of T47D and T47D-r to  $E_2$  and increasing concentrations of the antiestrogen ZM 182780 was measured with the MTT assay (Fig. 2). Cells were seeded in

96 well plates and immediately treated with ethanol, or  $10^{-10}$  M  $E_2$ , or  $10^{-10}$  M  $E_2$  plus ZM 182789 at concentrations from  $10^{-10}$  M to  $10^{-6}$  M. Whereas T47D are growth-inhibited by increasing concentrations of ZM 182780, the antiestrogen-resistant T47D-r grow equally well in medium with vehicle ethanol or increasing concentrations of ZM 182780.

**Differentially expressed genes** – RNA from T47D and T47D-r was hybridized to the Affymetrix arrays HuGeneFL and Hu95A. The results from Affymetrix DNA chip hybridization are displayed in Table 1. Out of a total of 6574 probe sets which interrogate genes and ESTs and which are represented on the HuGeneFL array, 104 genes were found consistently up-, and 63 down-regulated more than 2-fold in two independent cell culture experiments in T47D-r versus T47D cells. Out of a total of 12387 probe sets on the Hu95A array, 231 genes were up-, 316 genes down-regulated when RNA from a single cell culture experiment was hybridized. To determine the number of genes which are interrogated more than once on the two different Affymetrix arrays we counted the number of probe sets which are represented as singletons or are members of a cluster resulting in a non-redundant gene data set. First, from the total of 18961 probe sets present on both chips, 91 probe sets had to be removed because they contained low complexity gene information. The remaining 18870 probe sets fall into two groups with 3777 singletons and 7142 clusters containing 15093 sequences which result in 10919 non-redundant genes (Table 1). The same type of analysis was performed for the up- and down-regulated probe sets on the two array types. Taken together, 266 non-redundant genes are up-regulated and 300 are down-regulated in T47D-r compared to T47D (Table 1). An in-depth analysis of the Affymetrix DNA chip hybridization results will be described elsewhere (Footnote 1).

**2D-PAGE and image analysis** – In order to obtain well distributed spot patterns on overview gels (pH 3–10), the lysis conditions for the human breast cancer cell line T47D and its derivative T47D-r were optimized. Reproducible protein expression patterns were observed with the majority of proteins visible between pH 5–7. To increase the resolution

of the gels within this pH range and to obtain a higher number of well separated spots, the samples from three independent cell culture experiments were analyzed on narrow pH gradients (pH 4–7 and pH 6–11) (Fig. 3). For image comparison, the gels were scanned and automatic spot detection was performed with the Phoretix-2D advanced software. Depending on the spot detection parameters, manual editing of the images was found to be necessary mainly to split spots, to erase false spots and to include missed spots. These efforts resulted in approximately 2500 spots on the gels. After matching the gels derived from the T47D cell line with the gels from its derivative T47D-r, many spots appeared to be variant by at least a factor of two in spot volume (Fig. 3 and 4). In the pH range 4–7 53 proteins and in the pH range 6–11 another 31 proteins were differentially expressed. Of these 84 proteins 46 were consistently up-, and 38 down-regulated in the three independent cell culture experiments of T47D-r versus T47D.

**Identification of regulated proteins** – Proteomics analysis identified 84 protein spots that were consistently elevated 2-fold or more in three independent cell culture experiments of T47D or T47D-r (Fig. 5). Of these, 73 differential protein spots were excised from preparative Coomassie-stained 2D-gels. As the amount of protein loaded on preparative Coomassie-stained gels is on the upper limit of the capacity of the IPG strips, we established a further pH gradient (pH 5–8) to get a good resolution in the overlap region of the pH gradients pH 4–7 and pH 6–11. The excised spots were subjected to trypsin digestion and resulting peptide fragments were analyzed by MS. Assignments have been obtained for 66 out of 73 spots. Among the 66 identified protein spots 11 proteins were found in more than one spot, e.g. cathepsin D was identified in six independent spots (Fig. 4, Table 2) and MxA in two independent spots (Table 2). As a consequence, 47 distinct proteins were identified. To confirm differential protein expression of cathepsin D a Western Blot analysis was performed (Fig. 6) and the blot was quantitated after normalization to  $\beta$ -actin. Cathepsin D is down-regulated 2.4-fold in T47D-r treated with  $E_2$  and ZM 182780 compared to T47D treated with  $E_2$ .

**Differentially expressed proteins** – The differentially regulated proteins belong to several functional categories (Table 2 and 3). Thirteen proteins are enzymes and for seven the function is unknown. Six signal transduction proteins and six proteins with a role in protein degradation were identified. Four proteins are of mitochondrial origin, two each have a function in the cytoskeleton, are secreted proteins or are involved in transformation. Of the 47 proteins, 44 were either up- or down-regulated, whereas three proteins (stathmin/oncoprotein 18, 14-3-3 gamma, and Nudix hydrolase NUDT5), which each were identified from two or three independent spots on the gel, were regulated both, up in one spot and down in another, in T47D-r compared to T47D (Fig. 5). Concerning the 44 proteins regulated in only one direction, 20 proteins were up-regulated and 24 proteins were down-regulated in the comparison of T47D-r versus the T47D.

**Comparison of proteomics and Affymetrix data** – Only the 44 distinct proteins which were regulated up or down were compared to mRNA expression data (Fig. 5). To achieve this, the sequence of the protein which was identified by MS or the corresponding DNA sequence was used to perform a tBLASTn or BLASTn search, respectively, of the databases which contain the genes present on the Affymetrix arrays. This analysis showed that of the 44 proteins six were not present on the two arrays. Comparison of differential mRNA with protein expression data revealed that 19 out of 38 proteins for which the corresponding genes were present on the two arrays were regulated in the same direction on the mRNA and protein level, i.e. 50% of the proteins show a concordant regulation on the mRNA and protein level.

However, the strength of differential expression can vary. For 11 proteins the corresponding mRNAs are not well defined, i.e. either only one differential expression value is available and the fold-change is below 2 or the fold change is weak and not consistent between the experiments. Eight proteins were inversely regulated on the protein compared to the mRNA level. The differential expression of these mRNAs was



often below 1.8-fold (Table 2) whereas in standard Affymetrix experiments only genes regulated by a factor of two and higher are considered for subsequent studies.

## **Discussion**

**Comparison of differential proteomics and RNA expression analysis** – So far only a limited number of studies compared differential gene and protein expression (38, 40-47). In these studies cancer and normal tissue, different cell lines, and treated and untreated cellular model systems (e.g. yeast) were compared. The correlation of differential RNA and protein expression varied between moderate and no correlation according to these reports. Poor correlation was seen in a study of human liver (39), whereas a moderate correlation was reported in yeast (40). One group reported that for well resolved and abundant known proteins a highly significant correlation between protein levels as judged by 2D-PAGE and gene expression profiling analysis in non-invasive and invasive human transitional cell carcinomas was found (38). Several mRNAs and proteins showed a striking correspondence although in some cases also discrepancies were seen that could potentially be attributed to mRNA stability, splicing, translational regulation, post-translational processing, control of protein turn-over, protein degradation or a combination of these.

To extend previous studies it was our aim to analyze differential gene expression on both the mRNA level by Affymetrix DNA chip hybridization and on the protein level by proteomics. In addition, we expected to achieve a better insight into the changes which accompany the development of antiestrogen-resistance in our system.

Affymetrix DNA chip hybridizations were carried out on the commercially available HuGeneFL and Hu95A array. In sum, a total of 10919 non-redundant genes and ESTs were analyzed on the two arrays. A total of 266 genes (2.4%) were found to be up-

regulated and 300 (2.7%) down-regulated more than 2-fold in T47D-r compared to T47D (Table 1).

Compared to the proteomics analysis, where only 84 differentially regulated spots were identified on 2D-gels, far more genes were found to be differentially expressed. This can be explained by mRNA measurements having on the one hand a higher degree of sensitivity and the capability to detect also weakly expressed mRNAs and, on the other hand, oligonucleotides arrayed in parallel allowing for a higher degree of automation and throughput.

When the total cellular protein content from three independent cell culture experiments of T47D and T47D-r was analyzed, 84 spots were found to be reproducibly variant more than 2-fold. 66 proteins were identified corresponding to 47 distinct proteins because 11 proteins were present in more than one spot on the gel. The fact that one protein is localized in two different spots on the gel can be explained by post-translational modifications, e.g. phosphorylation, glycosylation or limited proteolytic cleavage (48, 49).

Three proteins (stathmin/oncoprotein 18, 14-3-3 gamma, and Nudix hydrolase NUDT5), were identified from two or three independent spots on the gel and were regulated both, up in one spot and down in another, in T47D-r compared to T47D (Fig. 5). The regulation in different directions in separate spots could also be the result of post-translational modifications. For instance, for 14-3-3 gamma and stathmin signal-induced phosphorylation has been described (50, 51). These three distinct proteins were excluded from the comparison with the Affymetrix results because their total amount was not investigated in this experimental setup. Another six proteins, which were not present on the two arrays hybridized, were also not included in the comparison with the Affymetrix results (Fig. 5).

The comparison of differential mRNA with protein expression data revealed that 19 out of 38 distinct proteins were up- or down-regulated in parallel at the mRNA and protein level, i.e. there is a concordance for the direction of changes on mRNA transcript and

protein expression level for 50% of the distinct proteins for which a probe set was present on the two arrays.

In 11 cases the differential expression level was not well-defined, therefore no comparison between protein and mRNA regulation could be made. For eight proteins an inverse regulation on protein versus mRNA level was observed. This inconsistency between differential mRNA transcript and protein expression levels in a steady-state system, i.e. antiestrogen-sensitive and antiestrogen-resistant human breast cancer cells, as investigated here could be due to the presence of regulatory influences which are currently not understood (52). All explanations taken into account would require that mechanisms as detailed below have to work differentially between the parental cell and its antiestrogen-resistant derivative. In case the protein is up-regulated and mRNA is unchanged or down-regulated one could postulate that either this particular protein is translated with higher efficiency, or it has a longer half-life, or that due to post-translational modification the protein appears *de novo* in an independent spot. When the protein is down-regulated while the mRNA is unchanged or up-regulated, the protein could have a shorter half-life, or it is translated with a lower efficiency. Further possibilities are that post-translational modification mechanism are turned off or that the protein is degraded into products no longer detectable because of various technical reasons (e.g. molecular weight or detection limit) leading to disappearance or decreasing intensity of this particular spot.

Since only those spots were considered to be differentially expressed in the proteomics analysis, that were regulated by a factor of two or more, it is conceivable that a given protein is present in further spots for reasons described above but was not detected as being regulated or expressed below the detection limit. If the identity of all further possible isoforms of a protein has not been analyzed completely, this may influence the correlation between mRNA and protein and represents a limitation. In addition, when one looks in detail at the fold regulation on the mRNA level these factors are often below 1.8. In a standard gene expression analysis only mRNAs with factors  $\geq 2$ -fold are considered

for subsequent analyses. A comparison of individual protein quantities from 2D gels with mRNA expression data has therefore to be performed with caution.

**Biological interpretation of the data** – In order to investigate the identified proteins further, we sorted them according to their functional category (Table 3). Several of them are housekeeping enzymes, mitochondrial proteins or involved in protein degradation. However, we also found proteins which play a role in signaling pathways or whose function is currently unknown. For several proteins which were identified in the proteomics analysis and by DNA chip hybridization, their role in regard to regulation by E<sub>2</sub> and endocrine resistance will be depicted.

The protease cathepsin D was down-regulated on the protein and mRNA level in T47D-r compared to T47D. The protein was found at six different locations at the same apparent molecular mass (33 kDa) but different isoelectric points (between pH 5 and 6.5) on the 2D-gels (Fig. 4). The different locations indicate that isoforms of cathepsin D exist. Cathepsin D is synthesized as pre-proenzyme that is post-translationally phosphorylated (53, 54). The 52 kDa pro-cathepsin D is cleaved to the enzymatically active 48 kDa heterodimer consisting of a 14 and 34 kDa subunit, and only the latter was identified in this proteomics study. The fold change on the protein level varied for the six spots between undetectable in T47D-r compared to T47D and down-regulated 22.9- to 2.1-fold (Table 2). In the Western Blot analysis the cathepsin D form of 32 kDa was down-regulated 2.4-fold in T47D-r compared to T47D.

On the mRNA level cathepsin D was down-regulated by a factor of more than 7.5- and 4.3-fold on the HuGeneFL array in T47D-r compared to T47D in the first and second cell culture experiment, respectively, and in the third experiment on the Hu95A array cathepsin D was down-regulated 7.3-fold.

Already in an early proteomics study on breast cancer cell lines a secreted 46 kDa protein was identified which was induced by E<sub>2</sub> in several human estrogen receptor positive cell lines and later determined by specific antibodies to be the protease cathepsin D (55).

Thus, the cathepsin D gene is a well known E<sub>2</sub>-regulated gene and induction by E<sub>2</sub> has been demonstrated for the ER-positive breast cancer cell lines T47D, ZR-75-1 and MCF-7 (56, 57). Since T47D-r cells have lost expression of ER alpha the E<sub>2</sub>-induction of cathepsin D is abrogated when compared to T47D cells. Differential expression of cathepsin D was found in both human breast cancer cell lines and breast biopsies and reduction mammoplasties

The small GTPase Rab11a was down-regulated in T47D-r compared to T47D on the protein and mRNA level. On the protein level Rab11a was found to be decreased 6.6-fold in T47D-r versus T47D. In the HuGeneFL and Hu95A array hybridization experiments Rab11a was down-regulated by a factor of 1.8, 2.8 and 2.3 in mRNA from three independent cell culture experiments. Rab11a belongs to the large family of Ras-related small GTP-binding proteins which are involved in the regulation of vesicular transport along the endocytotic and exocytotic pathway (58, 59). It has been shown that Rab11a is regulated by E<sub>2</sub> in the human endometrial cell line Ishikawa after transfection with either ER alpha or beta, whereas treatment with ZM 182780 inhibited the E<sub>2</sub>-dependent induction of Rab11a (60). The fact that Rab11a is down-regulated on the protein and mRNA level in the breast carcinoma cell line T47D-r, which is continuously treated with ZM 182780 and has lost ER alpha expression, indicates that also in this cell line expression of Rab11a might at least be partially under the control of the ER. Interestingly, the Rab11a protein was recently found to be up-regulated in matched pairs of microdissected ductal carcinoma *in situ* (DCIS) compared to normal ductal breast tissue by an extensive 2D-PAGE-based proteomics analysis and results were confirmed by immunohistochemical staining (32).

Another GTPase, the MxA protein, was identified from two spots in T47D-r. Comparing the spot intensities, the up-regulation was found to be 2- and 5.6-fold in T47D-r versus T47D, respectively. On the mRNA level MxA was strongly up-regulated by more than 8.7-, 5.2- and 3.6-fold in T47D-r versus T47D in the first, second and third chip

hybridization experiment on the HuGeneFL and Hu95A arrays, respectively. MxA, which belongs to the dynamin family of GTPases (61), is induced by interferon alpha and beta and the promoter contains two functional interferon response elements (ISREs) (62-64). One could assume that the transcription factors STAT1 and STAT2 (signal transducer and activator of transcription), which bind to ISREs (65) are involved in the up-regulation of MxA. Thus the up-regulation seen for MxA in T47D-r can potentially be explained by an activated STAT pathway in these cells.

hAG-2, which is the human homologue of the *Xenopus laevis* anterior gradient 2 protein (gene symbol AGR2), was down-regulated on the protein and mRNA level in T47D-r. The fold regulation on the mRNA level was 5.3-fold on the Hu95A array. On the protein level hAG-2 is decreased by a factor of 2.9 in T47D-r compared to T47D. On the mRNA level, hAG-2 has been shown to be co-expressed with the ER in breast cancer cell lines (66). As the T47D-r cells have lost protein and mRNA expression of the ER alpha (8), the identification of hAG-2 protein in the ER-positive T47D cells confirms the results shown previously on the mRNA level. Although the function of hAG-2 in the mammalian system is unknown, based on the expression data it was proposed that hAG-2 might be involved in the tumor cell biology of hormone-responsive, well differentiated breast cancer (66).

In a proteomics study in which the plasma membranes of several ER alpha and EGFR-positive human breast cancer cell lines were enriched the hAG-2 homologous protein hAG-3 was identified (34) and in a focussed follow-up study hAG-2 and hAG-3 were analyzed by immunohistochemical methods (67). Here, it was demonstrated by quantitative RT-PCR and immunohistochemistry on tissue microarrays that hAG-2 expression correlated positively with estrogen receptor and negatively with EGFR expression (67). Bioinformatic analysis indicated that the hAG-2 promoter contains four putative estrogen response elements. In a yeast 2-hybrid screen hAG-2 was shown to interact with alpha-dystroglycan (DAG-1) and GPI-anchored protein C4.4a (67) and it was suggested that hAG-2 as secreted protein interacts with the extracellular matrix.

Many other proteins were found to be variant in our model system, which have not yet been described in association with the development of breast cancer and which are now further investigated to understand their role in endocrine resistant breast cancer.

**Conclusion** – In the present study we identified and compared differentially expressed genes and proteins in a model system of antiestrogen-sensitive and -resistant breast cancer and evaluated consistency between mRNA and protein data. In our steady-state model concordance was 50%. By combining proteomics and Affymetrix technologies we achieved a more detailed insight into the expression changes that accompany the development of antiestrogen-resistance in an in vitro system. RNA expression profiling on Affymetrix oligonucleotide arrays is a complementary method to proteomics as gel-based proteomics is often unable to detect the low-abundance proteins. Several known E<sub>2</sub>-regulated genes like cathepsin D or genes co-expressed with the estrogen receptor alpha and likely also E<sub>2</sub>-regulated like hAG-2 are expressed in the parental T47D cell line. As manifestation of the antiestrogen-resistant phenotype of the T47D-r cells cathepsin D and hAG-2 are down-regulated on the RNA and protein level. The discovery of gene and protein expression changes in the context of antiestrogen resistance may allow for a better understanding of the mechanisms which lead to endocrine-resistant breast carcinoma.

## References

1. Parczyk, K., Schneider, M. R. (1996). The future of antihormone therapy: innovations based on an established principle. *J. Cancer Res. Clin. Oncol.* **122**, 383–396
2. Morris, C., Wakeling, A. (2002). Fulvestrant ('Faslodex') – a new treatment option for patients progressing on prior endocrine therapy. *Endocr. Relat. Cancer* **9**, 267–276
3. Howell, A., DeFriend, D., Robertson, J., Blamey, R., Walton, P. (1995). Response to a specific antioestrogen (ICI 182780) in tamoxifen-resistant breast cancer. *Lancet* **345**, 29–30
4. Howell, A., Robertson, J. F., Quaresma Albano, J., Aschermannova, A., Mauriac, L., Kleeberg, U. R., Vergote, I., Webster, A., Morris, C. (2002). Fulvestrant, formerly ICI 182,780, is as effective as anastrozole in postmenopausal women with advanced breast cancer progressing after prior endocrine treatment. *J. Clin. Oncol.* **20**, 3396–3403
5. Gottardis, M. M., Jiang, S. Y., Jeng, M. H., Jordan, V. C. (1989). Inhibition of tamoxifen-stimulated growth of an MCF-7 tumor variant in athymic mice by novel steroidal antioestrogens. *Cancer Res.* **49**, 4090–4093
6. Osborne, C. K., Coronado-Heinsohn, E. B., Hilsenbeck, S. G., McCue, B. L., Wakeling, A. E., McClelland, R. A., Manning, D. L., Nicholson, R. I. (1995). Comparison of the effects of a pure steroidal antiestrogen with those of tamoxifen in a model of human breast cancer. *J. Natl. Cancer Inst.* **87**, 746–750
7. Clarke, R., Skaar, T. C., Bouker, K. B., Davis, N., Lee, Y. R., Welch, J. N., Leonessa, F. (2001). Molecular and pharmacological aspects of antiestrogen resistance. *J. Steroid Biochem. Mol. Biol.* **76**, 71–84



8. Sommer, A., Hoffmann, J., Lichtner, R., Schneider, M., Parczyk, K. (2003). Studies on the development of resistance to the pure antiestrogen Faslodex<sup>TM</sup> in three human breast cancer cell lines. *J. Steroid Biochem. Mol. Biol.* **85**, 33-47
9. Brown, P. O., Botstein, D. (1999). Exploring the new world of the genome with DNA microarrays. *Nat. Genet.* **21** (1 suppl.), 33-37
10. Perou, C. M., Jeffrey, S. S., van de Rijn, M., Rees, C. A., Eisen, M. B., Ross, D. T., Pergamenschikov, A., Williams, C. F., Zhu, S. X., Lee, J. C., Lashkari, D., Shalon, D., Brown, P. O., Botstein, D. (1999). Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 9212-9217
11. Ross, D. T., Scherf, U., Eisen, M. B., Perou, C. M., Rees, C., Spellman, P., Iyer, V., Jeffrey, S. S., Van de Rijn, M., Waltham, M., Pergamenschikov, A., Lee, J. C., Lashkari, D., Shalon, D., Myers, T. G., Weinstein, J. N., Botstein, D., Brown, P. O. (2000). Systematic variation in gene expression patterns in human cancer cell lines. *Nat. Genet.* **24**, 227-235
12. Charpentier, A. H., Bednarek, A. I. C., Daniel, R. L., Hawkins, K. A., Laflin, K. J., Gaddis, S., MacLeod, M. C., Aldaz, C. M. (2000). Effects of estrogen on global gene expression: Identification of novel targets of estrogen action. *Cancer Res.* **60**, 5977-5983
13. Brenton, J. D., Aparicio, S. A., Caldas, C. (2001). Molecular profiling of breast cancer: portraits but not physiognomy. *Breast Cancer Res.* **3**, 77-80
14. Gruvberger, S., Ringner, M., Chen, Y., Panavally, S., Saal, L. H., Borg, A., Ferno, M., Peterson, C., Meltzer, P. S. (2001). Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns. *Cancer Res.* **61**, 5979-5984

15. Sørlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Thorsen, T., Quist, H., Matese, J. C., Brown, P. O., Botstein, D., Eystein Lonning, P., Borresen-Dale, A. L. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10869-10874
16. Hondermarck, H., Vercoutter-Edouart, A. S., Revillion, F., Lemoine, J., El-Yazidi-Belkoura, I., Nurcombe, V., Peyrat, J. P. (2001). Proteomics of breast cancer for marker discovery and signal pathway profiling. *Proteomics* **1**, 1216-1232
17. Hondermarck, H., Dolle, L., El Yazidi-Belkoura, I., Vercoutter-Edouart, A. S., Adriaenssens, E., Lemoine, J. (2002). Functional proteomics of breast cancer for signal pathway profiling and target discovery. *J. Mammary Gland Biol. Neoplasia* **7**, 395-405
18. Adriaenssens, E., Lemoine, J., El Yazidi-Belkoura, I., Hondermarck, H. (2002). Growth signaling in breast cancer cells: outcomes and promises of proteomics. *Biochem. Pharmacol.* **64**, 797-803
19. Yazidi-Belkoura, I. E., Adriaenssens, E., Vercoutter-Edouart, A. S., Lemoine, J., Nurcombe, V., Hondermarck, H. (2002). Proteomics of breast cancer: outcomes and prospects. *Technol. Cancer Res. Treat.* **1**, 287-296
20. Hondermarck, H. (2003). Breast cancer: when proteomics challenges biological complexity. *Mol. Cell. Proteomics* **2**, 281-291
21. Celis, J. E., Gromov, P., Gromova, I., Moreira, J. M., Cabezón, T., Ambartsumian, N., Grigorian, M., Lukanidin, E., Thor Straten, P., Guldborg, P., Bartkova, J., Bartek, J., Lukas, J., Lukas, C., Lykkesfeldt, A., Jäättelä, M., Roepstorff, P., Bolund, L., Ørntoft, T.,

Brüner, N., Overgaard, J., Sandelin, K., Blichert-Toft, M., Mouridsen, H., Rank, F. E. (2003). Integrating Proteomic and Functional Genomic Technologies in Discovery-driven Translational Breast Cancer Research. *Mol. Cell. Proteomics* **2**, 369-377

22. Giometti, C. S., Williams, K., Tollaksen, S. L. (1997). A two-dimensional electrophoresis database of human breast epithelial cell proteins. *Electrophoresis* **18**, 573-581

23. Williams, K., Chubb, C., Huberman, E., Giometti, C. S. (1998). Analysis of differential protein expression in normal and neoplastic human breast epithelial cell lines. *Electrophoresis* **19**, 333-343

24. Rasmussen, R. K., Ji, H., Eddes, J. S., Zugaro, L. M., Reid, G. E., Simpson, R. J., Dorow, D. S. (1998). Two-dimensional gel database of human breast carcinoma cell expressed proteins: an update. *Electrophoresis* **19**, 818-825

25. Gharbi, S., Gaffney, P., Yang, A., Zvelebil, M. J., Cramer, R., Waterfield, M. D., Timms, J. F. (2002). Evaluation of two-dimensional differential gel electrophoresis for proteomic expression analysis of a model breast cancer cell system. *Mol. Cell. Proteomics* **1**, 91-98

26. Pucci-Minafra, I., Fontana, S., Cancemi, P., Alaimo, G., Minafra, S. (2002). Proteomic patterns of cultured breast cancer cells and epithelial mammary cells. *Ann. N. Y. Acad. Sci.* **963**, 122-139

27. Pucci-Minafra, I., Fontana, S., Cancemi, P., Basirico, L., Caricato, S., Minafra, S. (2002). A contribution to breast cancer cell proteomics: detection of new sequences. *Proteomics* **2**, 919-927

28. Franzen, B., Linder, S., Alaiya, A. A., Eriksson, E., Uruy, K., Hirano, T., Okuzawa, K., Auer, G. (1996). Analysis of polypeptide expression in benign and malignant human breast lesions: down-regulation of cytokeratins. *Br. J. Cancer*. **74**,1632-1638
29. Franzen, B., Linder, S., Alaiya, A. A., Eriksson, E., Fujioka, K., Bergman, A. C., Jornvall, H., Auer, G. (1997). Analysis of polypeptide expression in benign and malignant human breast lesions. *Electrophoresis* **18**, 582-587
30. Bini, L., Magi, B., Marzocchi, B., Arcuri, F., Tripodi, S., Cintorino, M., Sanchez, J. C., Frutiger, S., Hughes, G., Pallini, V., Hochstrasser, D. F., Tosi, P. (1997). Protein expression profiles in human breast ductal carcinoma and histologically normal tissue. *Electrophoresis* **18**, 2832-2841
31. Bergman, A. C., Benjamin, T., Alaiya, A., Waltham, M., Sakaguchi, K., Franzen, B., Linder, S., Bergman, T., Auer, G., Appella, E., Wirth, P. J., Jornvall, H. (2000). Identification of gel-separated tumor marker proteins by mass spectrometry. *Electrophoresis* **21**, 679-686
32. Wulfkuhle, J. D., Sgroi, D. C., Krutzsch, H., McLean, K., McGarvey, K., Knowlton, M., Chen, S., Shu, H., Sahin, A., Kurek, R., Wallwiener, D., Merino, M. J., Petricoin, E. F., III, Zhao, Y., Steeg, P. S. (2002). Proteomics of human breast ductal carcinoma in situ. *Cancer Res.* **62**, 6740–6749
33. Page, M. J., Amess, B., Townsend, R. R., Parekh, R., Herath, A., Brusten, L., Zvelebil, M. J., Stein, R. C., Waterfield, M. D., Davies, S. C., O'Hare, M. J. (1999). Proteomic definition of normal human luminal and myoepithelial breast cells purified from reduction mammoplasties. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 12589-12594

34. Adam, P. J., Boyd, R., Tyson, K. L., Fletcher, G. C., Stamps, A., Hudson, L., Poyser, H. R., Redpath, N., Griffiths, M., Steers, G., Harris, A. L., Patel, S., Berry, J., Loader, J. A., Townsend, R. R., Daviet, L., Legrain, P., Parekh, R., Terrett, J. A. (2003). Comprehensive proteomic analysis of breast cancer cell membranes reveals unique proteins with potential roles in clinical cancer. *J. Biol. Chem.* **278**, 6482-6489
35. Chen, S. T., Pan, T. L., Tsai, Y. C., Huang, C. M. (2002). Proteomics reveals protein profile changes in doxorubicin-treated MCF-7 human breast cancer cells. *Cancer Lett.* **181**, 95-107
36. Celis, J. E., Kruhoffer, M., Gromova, I., Frederiksen, C., Ostergaard, M., Thykjaer, T., Gromov, P., Yu, J., Palsdottir, H., Magnusson, N., Orntoft, T. F. (2000). Gene expression profiling: monitoring transcription and translation products using DNA microarrays and proteomics. *FEBS Lett.* **480**, 2-16
37. Gygi, S. P., Rochon, Y., Franza, B. R., Aebersold, R. (1999). Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* **19**, 1720-1730
38. Ørntoft, T. F., Thykjaer, T., Waldman, F. M., Wolf, H., Celis, J. E. (2002). Genome-wide study of gene copy numbers, transcripts, and protein levels in pairs of non-invasive and invasive human transitional cell carcinomas. *Mol. Cell. Proteomics.* **1**, 37-45
39. Anderson, L., Seilhamer, J. (1997). A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* **18**, 533-537
40. Ideker, T., Thorsson, V., Ranish, J. A., Christmas, R., Buhler, J., Eng, J. K., Bumgarner, R., Goodlett, D. R., Aebersold, R., Hood, L. (2001) Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* **292**, 929-934

41. Le Naour, F., Hohenkirk, L., Grolleau, A., Misek, D. E., Lescure, P., Geiger, J. D., Hanash, S., Beretta, L. (2001) Profiling changes in gene expression during differentiation and maturation of monocyte-derived dendritic cells using both oligonucleotide microarrays and proteomics. *J. Biol. Chem.* **276**, 17920-17931
42. Chen, G., Gharib, T. G., Huang, C. C., Thomas, D. G., Shedden, K. A., Taylor, J. M., Kardia, S. L., Misek, D. E., Giordano, T. J., Iannettoni, M. D., Orringer, M. B., Hanash, S. M., Beer, D. G. (2002) Proteomic analysis of lung adenocarcinoma: identification of a highly expressed set of proteins in tumors. *Clin. Cancer Res.* **8**, 2298-2305
43. Chen, G., Gharib, T. G., Huang, C. C., Taylor, J. M., Misek, D. E., Kardia, S. L., Giordano, T. J., Iannettoni, M. D., Orringer, M. B., Hanash, S. M., Beer, D. G. (2002) Discordant protein and mRNA expression in lung adenocarcinomas. *Mol. Cell. Proteomics* **1**, 304-313
44. Griffin, T. J., Gygi, S. P., Ideker, T., Rist, B., Eng, J., Hood, L., Aebersold, R. (2002) Complementary profiling of gene expression at the transcriptome and proteome levels in *Saccharomyces cerevisiae*. *Mol. Cell. Proteomics* **1**, 323-333
45. Hanash, S. M., Bobek, M. P., Rickman, D. S., Williams, T., Rouillard, J. M., Kuick, R., Puravs, E. (2002). Integrating cancer genomics and proteomics in the post-genome era. *Proteomics* **2**, 69-75
46. Fessler, M. B., Malcolm, K. C., Duncan, M. W., Worthen, G. S. (2002). A genomic and proteomic analysis of activation of the human neutrophil by lipopolysaccharide and its mediation by p38 mitogen-activated protein kinase. *J. Biol. Chem.* **277**, 31291-31302

47. Washburn, M. P., Koller, A., Oshiro, G., Ulaszek, R.R., Plouffe, D., Deciu, C., Winzeler, E., Yates, J. R., III. (2003) Protein pathway and complex clustering of correlated mRNA and protein expression analyses in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 3107-3112
48. Imam-Sghiouar, N., Laude-Lemaire, I., Labas, V., Pflieger, D., Le Caer, J. P., Caron, M., Nabias, D. K., Joubert-Caron, R. (2002). Subproteomics analysis of phosphorylated proteins: application to the study of B-lymphoblasts from a patient with Scott syndrome. *Proteomics* **2**, 828–838
49. Djelloul, S., Tarunina, M., Barnouin, K., Mackay, A., Jat, P. S. (2002). Differential protein expression, DNA binding and interaction with SV40 large tumour antigen implicate the p63 family of proteins in replicative senescence. *Oncogene* **21**, 981–989
50. Autieri, M. V., Carbone, C. J. (1999). 14-3-3-gamma interacts with and is phosphorylated by multiple protein kinase C isoforms in PDGF-stimulated human vascular smooth muscle cells. *DNA Cell Biol.* **18**, 555–564
51. Brattsand, G. (2000). Correlation of oncoprotein 18/stathmin expression in human breast cancer with established prognostic factors. *Br. J. Cancer* **83**, 311–318
52. Pradet-Balade, B., Boulme, F., Beug, H., Mullner, E. W., Garcia-Sanz, J.A. (2001) Translation control: bridging the gap between genomics and proteomics? *Trends Biochem. Sci.* **26**, 225-229
53. Cuozzo, J. W., Toa, K., Cygler, M., Mort, J. S., Sahagian, G. G. (1998). Lysine-based structure responsible for selective mannose phosphorylation of cathepsin D and cathepsin L defines a common structural motif for lysosomal enzyme targeting. *J. Biol. Chem.* **273**, 21067–21076

54. Fioravanti, L., Cappelletti, V., Miodini, P., Ronchi, E., Brivio, M., Di Fronzo, G. (1998). Genistein in the control of breast cancer cell growth: insights into the mechanism of action in vitro. *Cancer Lett.* **130**, 143–152
55. Westley, B., Rochefort, H. (1980). A secreted glycoprotein induced by estrogen in human breast cancer cell lines. *Cell* **20**, 353-362
56. Westley, B. R., May, F. E. (1987). Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive human breast cancer cells. *Nucl. Acids Res.* **15**, 3773–3786
57. May, F. E., Westley, B. R. (1995). Estrogen regulated messenger RNAs in human breast cancer cells. *Biomed. Pharmacother.* **49**, 400–414
58. Zhu, A. X., Zhao, Y., Flier, J. S. (1994). Molecular cloning of two small GTP-binding proteins from human skeletal muscle. *Biochem. Biophys. Res. Commun.* **205**, 1875–1882
59. Gromov, P. S., Celis, J. E., Hansen, C., Tommerup, N., Gromova, I., Madsen, P. (1998). Human rab11a: transcription, chromosome mapping and effect on the expression level of host GTP-binding proteins. *FEBS Lett.* **429**, 359–364
60. Chen, D., Ganapathy, P., Zhu, L.-J., Xu, X., Li, Q., Bagchi, I. C., Bagchi, M. K. (1999). Potential regulation of membrane trafficking by estrogen receptor  $\alpha$  via induction of Rab11 in uterine glands during implantation. *Mol. Endocrinol.* **13**, 993–1004
61. van der Blik, A. M. (1999). Functional diversity in the dynamin family. *Trends Cell. Biol.* **9**, 96–102



62. Horisberger, M. A. (1992). Interferon-induced human protein MxA is a GTPase which binds transiently to cellular proteins. *J. Virol.* **66**, 4705–4709
63. Nakade, K., Handa, H., Nagata, K. (1997). Promoter structure of the MxA gene that confers resistance to influenza virus. *FEBS Lett.* **418**, 315–318
64. Ronni, T., Matikainen, S., Lehtonen, A., Palvimo, J., Dellis, J., Van Eylen, F., Goetschy, J. F., Horisberger, M., Content, J., Julkunen, I. (1998). The proximal interferon-stimulated response elements are essential for interferon responsiveness: a promoter analysis of the antiviral MxA gene. *J. Interferon Cytokine Res.* **18**, 773–781
65. Levy, D. E., Darnell, J. E. Jr. (2002). Stats: transcriptional control and biological impact. *Nat. Rev. Mol. Cell Biol.* **3**, 651–662
66. Thompson, D. A., Weigel, R. J. (1998). hAG-2, the human homologue of *Xenopus laevis* anterior gland gene XAG-2, is coexpressed with estrogen receptor in breast cancer cell lines. *Biochem. Biophys. Res. Commun.* **251**, 111–116
67. Fletcher, G. C., Patel, S., Tyson, K., Adam, P. J., Schenker, M., Loader, J. A., Daviet, L., Legrain, P., Parekh, R., Harris, A. L., Terrett, J.A. (2003). hAG-2 and hAG-3, human homologues of genes involved in differentiation, are associated with oestrogen receptor-positive breast tumours and interact with metastasis gene C4.4a and dystroglycan. *Br. J. Cancer* **88**, 579-585

## Footnotes

1 Sommer A., personal communication.

## Figure Legends

Fig. 1: Phase contrast microscopy of T47D and T47D-r. The photography was taken 3 days after seeding T47D cells in standard medium (phenol red-free RPMI 1640 medium supplemented with 10% FCS, 2 mM glutamine, 200 mU/ml insulin and  $10^{-10}$  M  $E_2$ ). The antiestrogen-resistant T47D-r cells were seeded in standard medium with  $10^{-6}$  M ZM 182780. Bar = 100  $\mu$ m.

Fig. 2: Response of T47D and T47D-r to  $E_2$  and the antiestrogen ZM 182780. Cell proliferation was measured with the MTT assay. Cells were seeded in 96 well plates and immediately treated with ethanol, or  $10^{-10}$  M  $E_2$ , or  $10^{-10}$  M  $E_2$  plus ZM 182789 at concentrations from  $10^{-10}$  M to  $10^{-6}$  M. Mitochondrial activity was measured on day 7. Proliferation activity for  $E_2$ -treated T47D and T47D-r, respectively, was set to 100. Whereas T47D are growth-inhibited by increasing concentrations of ZM 182780, the antiestrogen-resistant T47D-r grow equally well in medium with vehicle ethanol or increasing concentrations of ZM 182780.

Fig. 3: Representative silver-stained 2D-maps of total cellular extracts from T47D (left) and T47D-r (right) cells. Upper row: pH gradient 4–7 focussed for 59 kVh on an IPGphor instrument. Lower row: pH gradient 6–11 focussed for 49 kVh on a Multiphor instrument. 100  $\mu$ g of total protein was loaded on the first dimension for analytical silver-stained gels. All protein spots which were found to be reproducibly variant in the comparison of T47D and T47D-r 2D-gels by at least a factor of two in all three independent experiments are indicated in red.

Fig. 4: Zoom-in region from the IPG strip pH 4-7 for T47D (upper) and T47D-r (lower). All protein spots which were found to be reproducibly variant in the comparison of T47D

and T47D-r 2D-gels by at least a factor of two in all three independent experiments are indicated in red. The spots representing cathepsin D, which is down-regulated in T47D-r compared to T47, are colored in blue.

Fig. 5: Flow chart of proteomics data. From 84 spots which were reproducibly variant by a factor of  $>2$  in all protein extracts derived from three independent cell culture experiments 73 were picked from Coomassie-stained preparative gels and analyzed by MALDI mass spectrometry. From these, 66 proteins were identified. As some proteins were present in more than one spot, the number was reduced to 47 distinct proteins. Only the 44 proteins which were either up- or down-regulated were compared with Affymetrix data. Proteins were categorized as being either regulated in the same or inverse direction on mRNA and protein level, as having no change on the mRNA level or as being not represented on the Affymetrix arrays HuGeneFL and Hu95A.

Fig. 6: Expression of cathepsin D is down-regulated in T47D-r. T47D and T47D-r were treated with the vehicle ethanol,  $10^{-10}$  M E<sub>2</sub>,  $10^{-7}$  M ZM 182780, or  $10^{-10}$  M E<sub>2</sub> and  $10^{-7}$  M ZM 182780 for 24 hrs. Cell extracts were prepared in lysis buffer, equal amounts of protein were separated on 10% Bis-Tris gels and transferred to nitrocellulose membrane. Cathepsin D was detected with the monoclonal  $\alpha$ -cathepsin D antibody Ab-1 from Oncogene Research. Detection of  $\beta$ -actin indicates equal protein content in all lanes.

Table 1: Overview of genes differentially expressed in T47D-r versus T47D analyzed on Affymetrix arrays HuGeneFL and Hu95A

Table 2: Proteomics data for all analyzed protein spots and comparison with corresponding mRNA expression data

Table 3: Functional categories of identified proteins

## **Acknowledgements**

We thank Anja Goede, Anke Kühnel, Dania Reipert, Daniela Schilling, and Elzbieta Wiecko for excellent technical assistance. We thank Bertram Weiss for bioinformatics support. We are grateful to Henrik Seidel for contributing to the gene expression analysis with dedicated software.

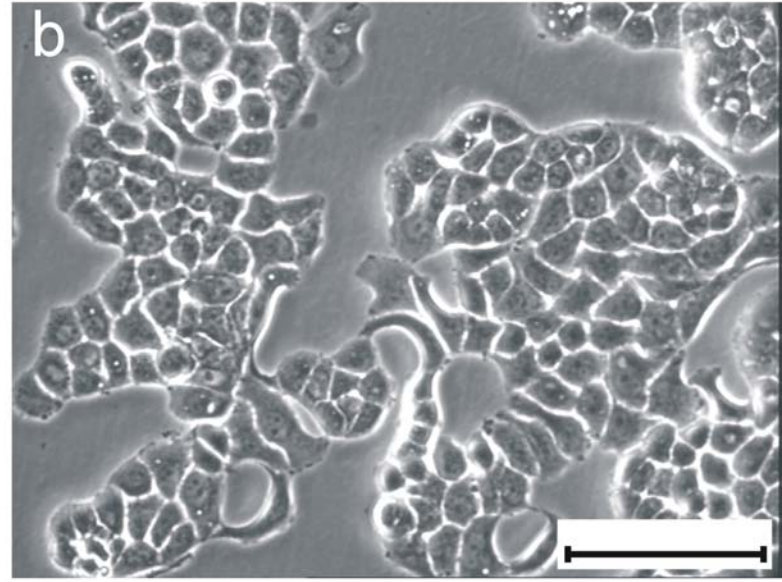
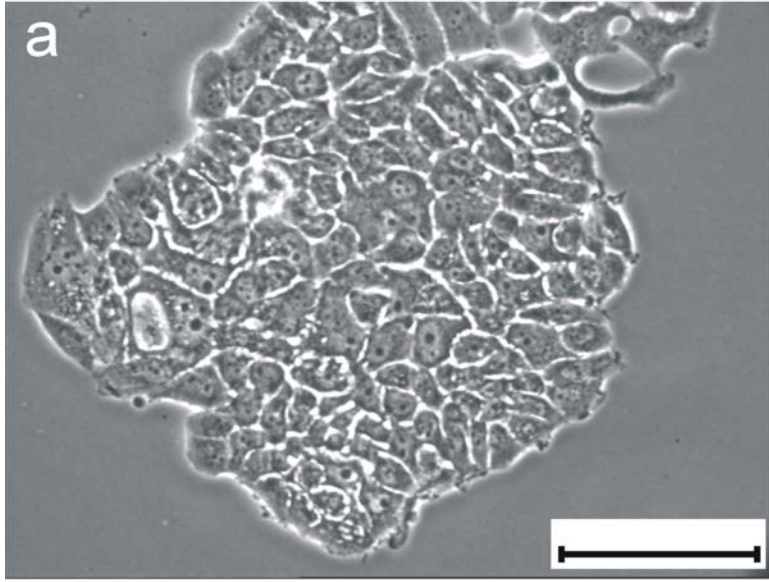


Fig. 1

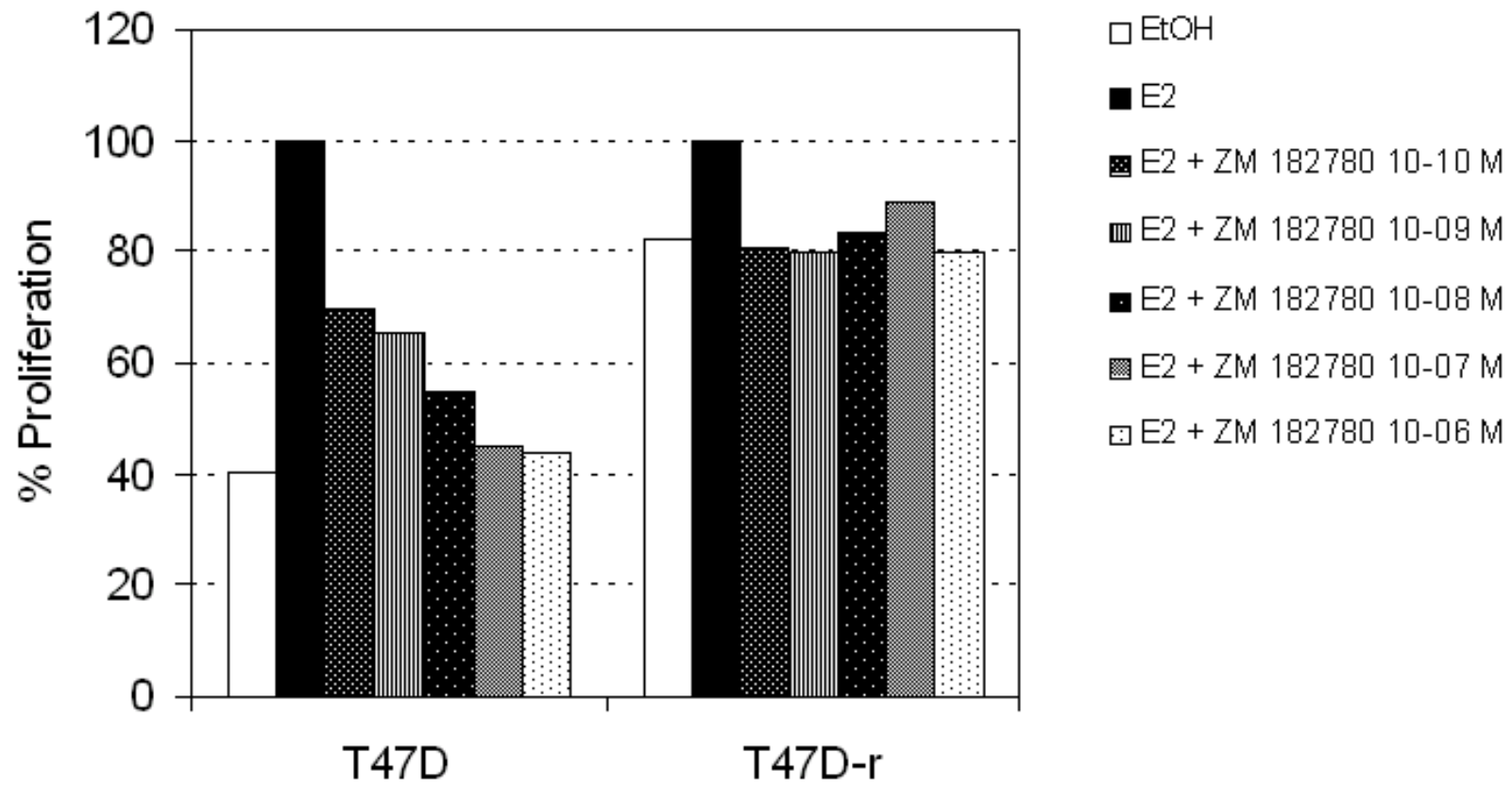


Fig. 2

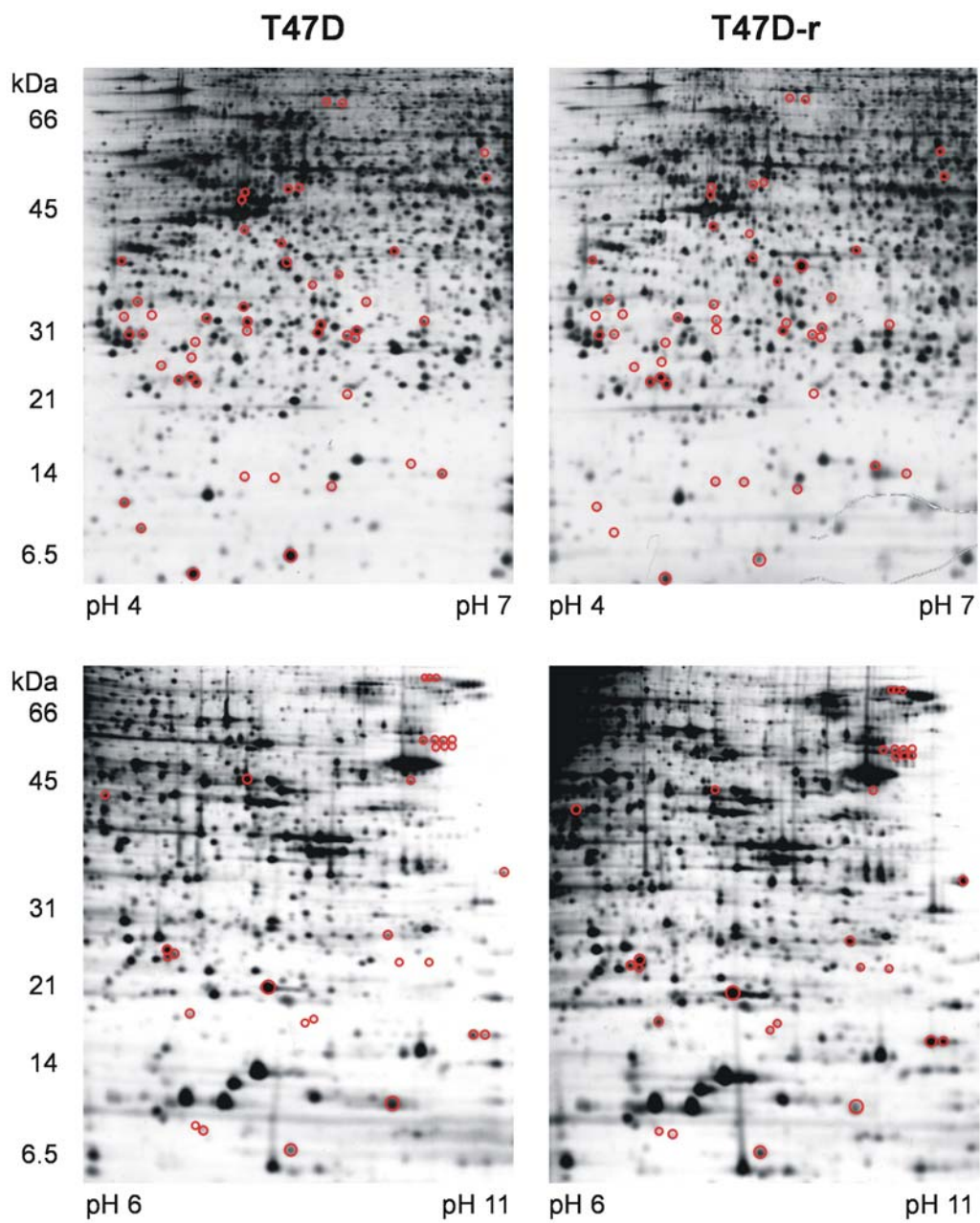
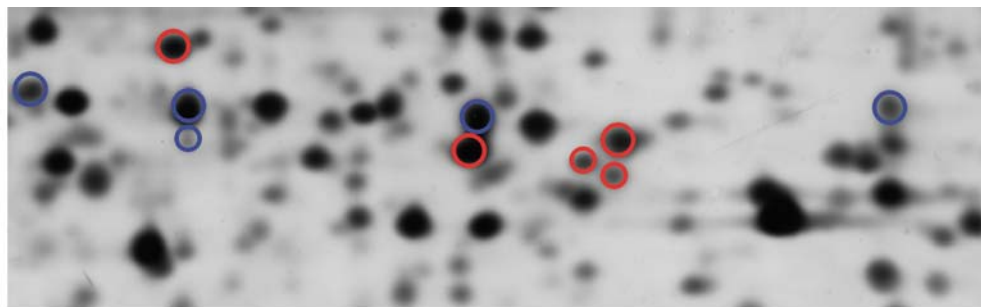
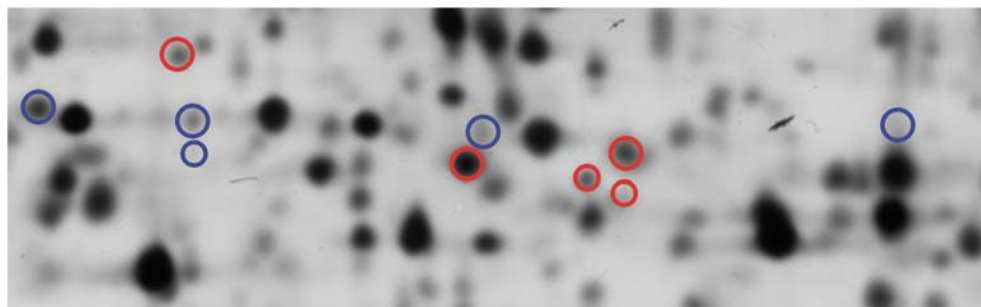


Fig. 3

T47D



T47D-r



pH 5

pH 6.5

Fig. 4





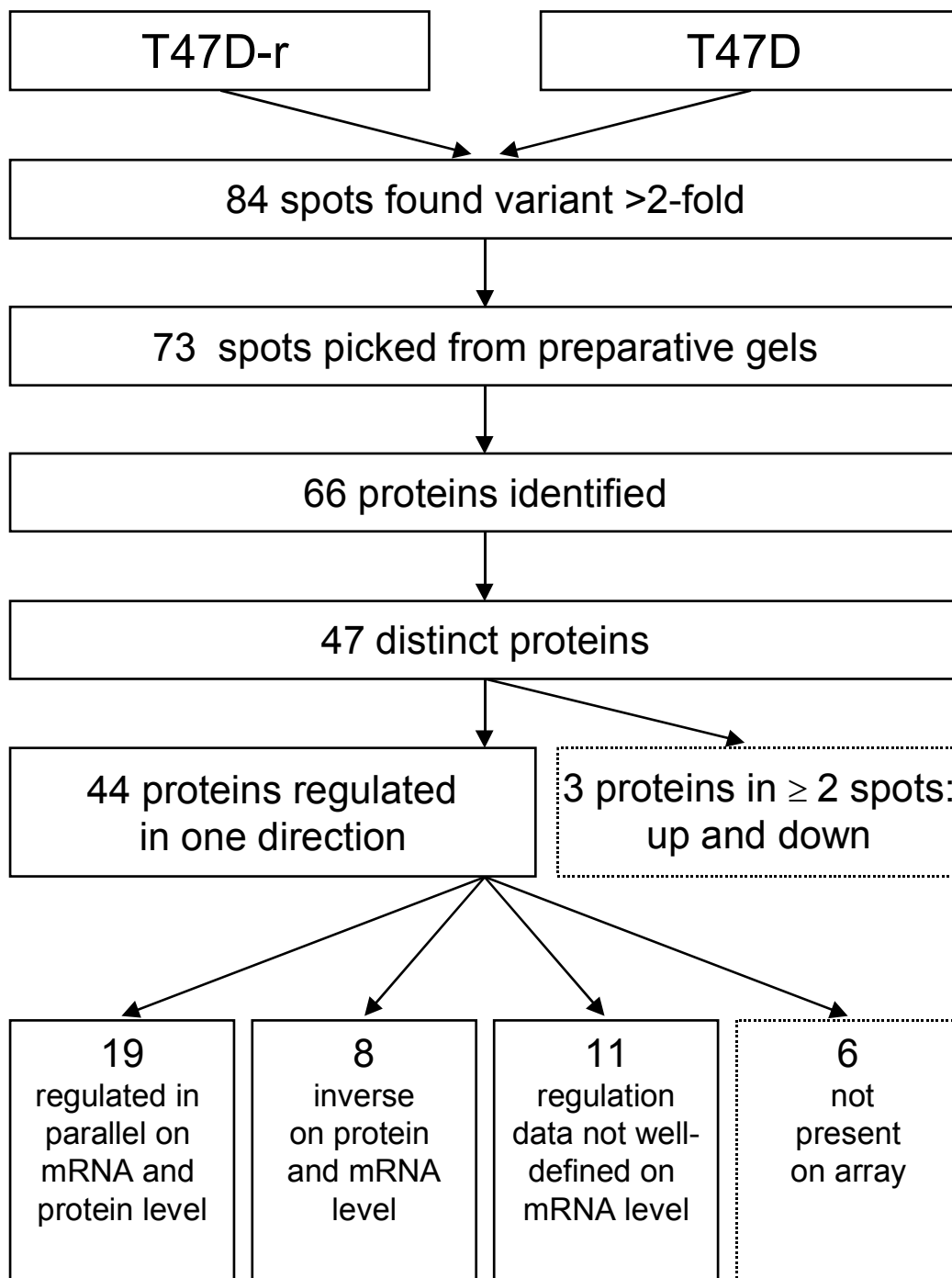


Fig. 5

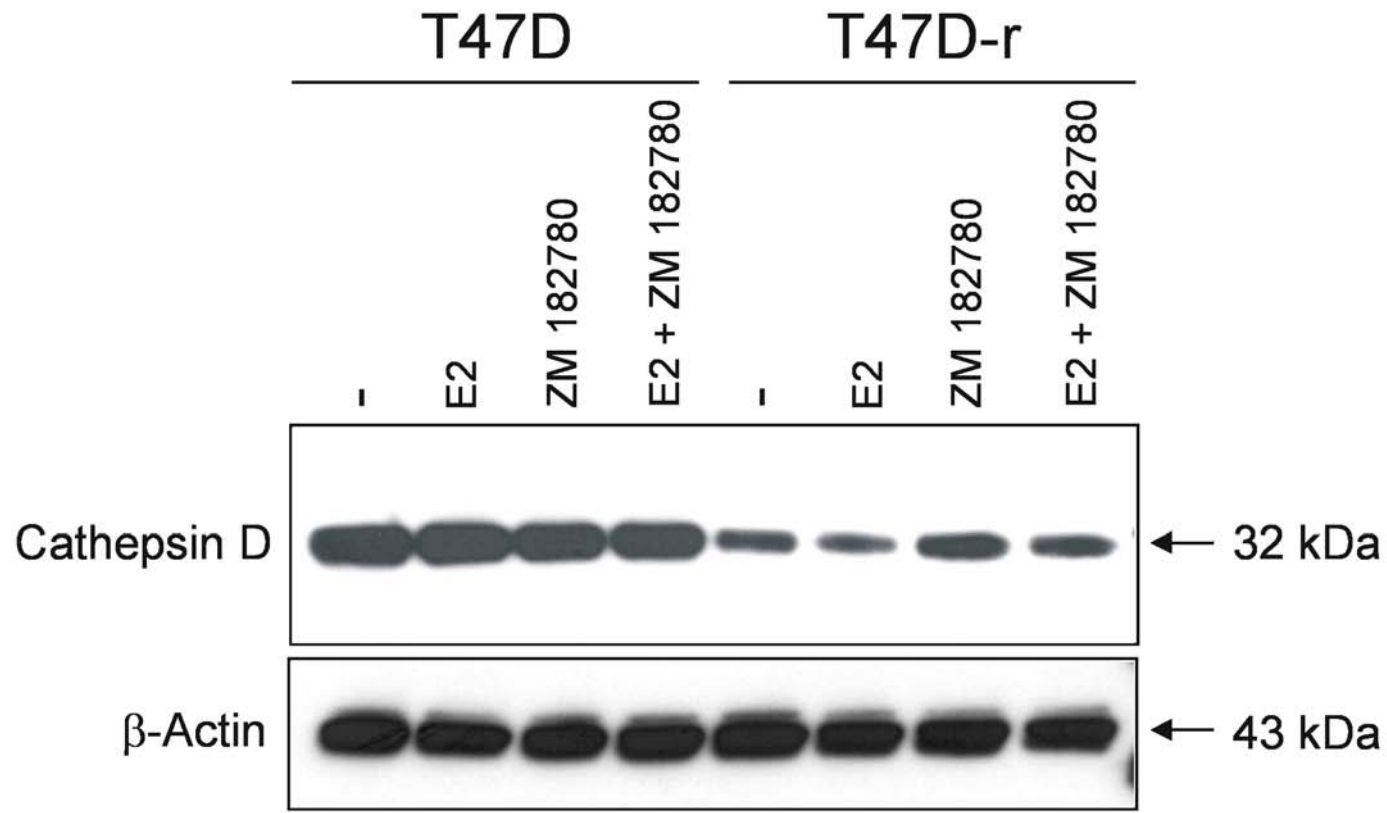


Fig. 6

**Table 1.** Overview of genes differentially expressed in T47D-r versus T47D analyzed on Affymetrix chips HuGeneFL and Hu95A

Array	Probe sets	up-regulated	down-regulated	total
HuGeneFL	6574	104	63	167
Hu95A	12387	231	316	547
HuGeneFL + Hu95A	18961	335	379	714
Non-redundant genes	10919	266	300	566

On the HuGeneFL array RNA from two independent cell culture experiments and on the Hu95A array RNA from one further independent cell culture experiment was hybridized. Genes which are regulated more than 2-fold on the arrays in T47D-r versus T47D are listed.

**Table 2** : Proteomics data for all analyzed protein spots and comparison with corresponding mRNA expression

Proteomics												DNA chip hybridization				
Spot ID	pH gradient	Fold regulation	GenBank protein identifier	Identified protein	Mr theor. (kDa)	Mr exp. (kDa)	pI theor.	pI exp.	matched pept./ total pept. in digest <sup>a</sup>	Error ppm <sup>b</sup>	seq. cov. %	Gene	Genbank nucleotide identifier	T47D-r versus T47D		
												Symbol	HuGeneFL 1. Exp.	HuGeneFL 2. Exp.	Hu95A	
<b>not present on chip</b>																
3	pH 4-7	6.3	gi 12652557 gi 13569956	similar to actin related protein 2/3 complex, subunit 5	17	15	6.2	6.3	4/5 <sup>c</sup>	9	33		NM_030978	np	np	np
6	pH 4-7	1.5	gi 4406577	unknow n, clone 24952	19	14	5.2	5.4	7/10 <sup>d</sup>	14	36		AF131758	np	np	np
33	pH 4-7	-2	gi 6018458	6-phosphogluconolactolase	28	25	5.7	5.9	11/12	30	49	PGLS	AJ243972	np	np	np
56	pH 6-11	only in T47D-r	gi 107356 gi 131528	polypyrimidine tract-binding protein PTB1 or PTB2	57	58	9.3	8.7	4/4	22	8	PTB2, PTB1	X66975	np	np	np
57	pH 6-11	1.8			57	58	9.3	8.7	5/6	18	10					
58	pH 6-11	2.1			57	58	9.3	8.7	6/6	11	12					
67	pH 5-8	-3.3	gi 6164955	vacuolar sorting protein VPS29/PEP11	21	25	6.3	7.2	5/5	18	32	VPS29	NM_016226	np	np	np
73	pH 5-8	3.9	gi 4680711	CGI-36 protein	19	27	6.4	7.2	9/11	20	33		NM_015963	np	np	np
<b>regulation data not well-defined on mRNA level</b>																
16	pH 4-7	-1.7	gi 3024348	platelet-activating factor acetylhydrolase	26	25	5.6	5.8	5/8	21	23	PAFAH1B2	D63390	np	np	1
26	pH 4-7	-2.3	gi 4759212	beta-tubulin cofactor	13	11	5.3	5.1	6/9	11	43	TBCA	AF038952	np	np	-1.5
35	pH 4-7	-2.1	gi 4506203	proteasome subunit	30	26	7.8	6.0	8/12	13	23	PSMB7	D38048	1.1	1.3	-1.1/-1.2
37	pH 4-7	-2.3	gi 4506003	protein phosphatase 1 alpha isoform	38	39	5.9	6.2	16/24	16	49	PPP1CA	S57501	np	np	-1.2
47	pH 6-11	2.9	gi 4505399	4-nitrophenylphosphatase domain and non-neuronal SNAP25-like 1	33	24	9.5	8.4	12/14	23	26	NIPSNAP1	AJ001258	np	np	-1.1

**Table 2 : continued**

Proteomics												DNA chip hybridization				
Spot ID	pH gradient	Fold regulation	GenBank protein identifier	Identified protein	Mr theor. (kDa)	Mr exp. (kDa)	pI theor.	pI exp.	matched pept./ total pept. in digest <sup>a</sup>	Error ppm <sup>b</sup>	seq. cov. %	Gene	Genbank nucleotide identifier	T47D-r versus T47D		
		T47D-r versus T47D										Symbol		HuGeneFL 1. Exp.	HuGeneFL 2. Exp.	Hu95A
48	pH 6-11	2.7	gi 4507791	ubiquitin-conjugating enzyme E2M	21	17	7.7	7.2	7/8	14	35	UBE2M	AF075599	np	np	1.8
49	pH 6-11	2.4 °	gi 4505591	Peroxiredoxin 1, proliferation-associated gene A (natural killer-enhancing factor A)	22	19	8.7	7.7	19/22	20	64	PRDX1	X67951	np	np	1
53	pH 6-11	2.2	gi 5454064	SYT interacting protein	70	80	9.7	8.7	16/24	16	26	RBM14	AF080561	np	np	-1.2
54	pH 6-11	3.8			70	80	9.7	8.7	10/12	22	17					
55	pH 6-11	2.1			70	79	9.7	8.7	9/12	12	16					
68	pH 5-8	2.3	gi 3859560	acyl-protein thioesterase	25	29	6.8	7.5	6/9	12	28	LYPLA2	NM_007260	np	np	1.3
72	pH 5-8	2.1	gi 509033	GARS protein	47	55	6.3	7.1	10/13	9	30	GARS	X54199	undef	1.1	1.2
74	pH 5-8	only in T47D-r	gi 547753	keratin, type II cytoskeletal 4	58	60	6.3	7.2	9/14	17	19	KRT4	X07695	undef	undef	1.3
<b>inverse regulation on protein and mRNA level</b>																
4	pH 4-7	-2.9	gi 238470 gi 251370	acid phosphatase isozyme Bf or Af	18	14	7.1	6.4	12/14	14	81	ACP1	U25849	1.4	1.7	-1.1
27	pH 4-7	-2	gi 4827038 gi 1706289	tumor protein D52	20	22	4.9	5.0	6/6 <sup>f</sup>	14	38	TPD52	U18914	undef	<1.3	1.2
34	pH 4-7	-6.2	gi 1352435	eukaryotic translation initiation factor 4E	25	25	5.8	6.0	6/7	31	22	EIF4E	M15353	1.1	-1.1	1.5
45	pH 6-11	-2.3	gi 4507777	ubiquitin-conjugating enzyme E2D 3	17	12	8.0	7.3	8/11	9	34	UBE2D3	U39318	2.2	1.2	1.5/1.3
60	pH 6-11	only in T47D-r	gi 4504327	hydroxyacyl-Coenzyme A dehydrogenase	52	57	9.6	8.7	9/10	25	21	HADHB	D16481	-1.2	-1.2	-1.4
61	pH 6-11	5.1			52	57	9.6	8.7	15/18	18	35					
62	pH 6-11	6.2			52	57	9.6	8.8	14/15	26	35					

**Table 2 : continued**

Proteomics												DNA chip hybridization				
Spot ID	pH gradient	Fold regulation	GenBank protein identifier	Identified protein	Mr theor. (kDa)	Mr exp. (kDa)	pI theor.	pI exp.	matched pept./ total pept. in digest <sup>a</sup>	Error ppm <sup>b</sup>	seq. cov. %	Gene	Genbank nucleotide identifier	T47D-r versus T47D		
		T47D-r versus T47D										Symbol		HuGeneFL 1. Exp.	HuGeneFL 2. Exp.	Hu95A
63	pH 6-11	1.6	gi 4758504	hydroxyacyl-Coenzyme A dehydrogenase, type II	27	22	8.0	7.1	15/17	20	75	HADH2	U73514	<-1.6	-1.5	np
64	pH 6-11	1.6			27	21	8.0	7.1	5/6	20	17					
65	pH 6-11	-2.5	gi 5902074	stromal cell-derived factor 2	23	21	7.2	7.1	6/7	31	35	SDF2	D50645	1.4	-1.1	-1
66	pH 5-8	-2	gi 4506195	proteasome subunit, beta type, 2	23	28	6.5	7.4	13/18	11	50	PSMB2	D26599	1.7	1.2	1.1
<b>regulated in parallel on mRNA and protein level</b>																
1	pH 4-7	-3.6	gi 999882	cellular retinoic acid-binding protein II	16	11	5.4	5.6	16/20	14	82	CRABP2	M97815	-2.5	-2.8	-1.6
7	pH 4-7	1.5	gi 2392338	glyoxalase I	21	20	5.1	5.0	10/14	14	51	GLO1	D13315	1.4	1.6	1.2
9	pH 4-7	1.6			21	20	5.1	5.1	9/14	20	39					
8	pH 4-7	1.5	gi 4502969	catechol-O-methyltransferase	30	21	5.3	5.1	12/19	18	47	COMT	Z26491	2.1	1.9	np
													M58525	>2.3	-1.1	2.2
12	pH 4-7	-9.3	gi 2281474	BCSG1 protein (synuclein-gamma, breast cancer-specific protein 1)	13	12	5.4	4.8	8/11	14	38	SNCG	AF010126	np	np	-2.2
13	pH 4-7	-11.8	gi 229383	cytochrome b5 fragment	10	13	5.2	4.8	5/6	5	75	CYB5	L39945	np	np	-2
													M22976	-1.3	-3.4	undef
14	pH 4-7	only in T47D	gi 4503143	cathepsin D	45	27	6.1	6.3	10/12	14	22	CTSD	M63138	<-7.5	-4.3	-7.3
17	pH 4-7	-22.9			45	26	6.1	5.8	11/14	6	26					
20	pH 4-7	only in T47D			45	26	6.1	5.4	6/7	20	13					
21	pH 4-7	-5.9			45	27	6.1	5.4	10/10	11	24					
32	pH 4-7	-2.1			45	27	6.1	5.2	5/7	28	8					
69	pH 5-8	only in T47D			45	36	6.1	6.8								

Table 2 : continued

Proteomics												DNA chip hybridization				
Spot ID	pH gradient	Fold regulation	GenBank protein identifier	Identified protein	Mr theor. (kDa)	Mr exp. (kDa)	pI theor.	pI exp.	matched pept./ total pept. in digest <sup>a</sup>	Error ppm <sup>b</sup>	seq. cov. %	Gene	Genbank nucleotide identifier	T47D-r versus T47D		
		T47D-r versus T47D										Symbol		HuGeneFL 1. Exp.	HuGeneFL 2. Exp.	Hu95A
15	pH 4-7	-6.6	gi 4758984	RAB 11A	24	19	6.1	5.9	5/8	3	21	RAB11A	AF000231	-1.8	-2.8	-2.3
18	pH 4-7	4.2	gi 1421662	annexin A3	36	32	5.6	5.7	16/18	9	41	ANXA3	L20591 M20560	3.7 np	1.7 np	np 7.5
19	pH 4-7	7	gi 4557032	lactate dehydrogenase B	37	36	5.7	5.9	10/11	13	27	LDHB	X13794	undef	>10	1.1
22	pH 4-7	-3.2	gi 15295949	hypothetical protein FLJ13612	27	29	5.3	5.4	12/14 <sup>f</sup>	19	59		NM_025202	np	np	-4.4
23	pH 4-7	only in T47D			27	27	5.3	4.7	4/4	14	17					
25	pH 4-7	-7.8 <sup>g</sup>	gi 227920	beta galactoside soluble lectin	15	7	5.3	5.0	8/12	12	54	LGALS1	J04456	-1.6	-1.6	-1.9
38	pH 4-7	-2	gi 387010	pyruvate dehydrogenase E1-beta subunit precursor	37	36	5.4	5.6	13/19	11	44	PDHB	D90086	1.1	-1.6	-1.9
39	pH 4-7	3.5	gi 4758904	serine protease inhibitor 6	43	44	5.2	5.4	12/18	12	32	SERPINB6	S69272	1.5	2.9	1.5
42	pH 4-7	5.6	gi 4505291	MxA/p78	76	94	5.6	5.9	23/25	11	32	MX1	M33882	>8.7	>5.2	>3.6
43	pH 4-7	2			76	94	5.6	5.8	20/22	11	29					
44	pH 6-11	-2.9	gi 5453541	anterior gradient 2	20	13	9.0	8.4	10/11	12	50	AGR2	AF038451	np	np	-5.3
46	pH 6-11	-19.1	gi 4557337	argininosuccinate synthetase	47	50	8.5	7.5	11/14	20	23	ASS	X01630	-1.3	-5	-3.1
52	pH 6-11	2.5	gi 2624694	Chain A, human mitochondrial single-stranded DNA binding protein	15	11	8.2	7.9	12/14	18	79	SSBP1	M94556	1.6	1.3	1.2
70	pH 5-8	-4.1	gi 4557587 gi 31291	fumarylacetoacetase	47	50	6.5	7.2	8/10	14	25	FAH	M55150	-1.9	-1.4	<-2
71	pH 5-8	only in T47D	gi 5729980	phosphomevalonate kinase	22	26	5.6	6.3	13/14	16	35	PMVK	L77213	-1.4	-1.7	-1.9



Table 2 : continued

Proteomics												DNA chip hybridization				
Spot ID	pH gradient	Fold regulation	GenBank protein identifier	Identified protein	Mr theor. (kDa)	Mr exp. (kDa)	pl theor.	pl exp.	matched pept./ total pept. in digest <sup>a</sup>	Error ppm <sup>b</sup>	seq. cov. %	Gene	Genbank nucleotide identifier	T47D-r versus T47D		
												Symbol	HuGeneFL 1. Exp.	HuGeneFL 2. Exp.	Hu95A	
<b>regulated up and down on protein level</b>																
2	pH4-7	-2.2	gi 5031851	stathmin	17	14	5.8	5.8	10/14 <sup>d</sup>	4	50	STMN1	M31303	1.4	1.4	1.4
5	pH4-7	3.6			17	14	5.8	5.5	4/6 <sup>d</sup>	9	25					
10	pH4-7	2.1	gi 6016838	14-3-3 protein gamma	28	25	4.8	4.8	11/16	5	45	YWHA3	AB024334	np	np	np
11	pH4-7	-2.6			28	25	4.8	4.8	10/13	9	36					
24	pH4-7	3	gi 6694937 gi 6288769	adenosine 5'-diphosphosugar pyrophosphatase	25	28	4.9	4.9	6/8	23	27	NUDT5	AF218818	np	np	np
28	pH4-7	2.4			25	30	4.9	4.8	7/8 <sup>f</sup>	15	29					
29	pH4-7	-2			25	37	4.9	4.7	4/4	25	14					

np: not present on array

undef: Undefined, indicating that in the two compared chip hybridizations no signal was detected.

<sup>a</sup> autolytic tryptic peptides and contaminants were not taken into account

<sup>b</sup> average error of matched peptide masses

<sup>c</sup> two peptides confirmed by ESI-MS/MS

<sup>d</sup> one peptide confirmed by MALDI-PSD

<sup>e</sup> spot saturated

<sup>f</sup> one peptide confirmed by ESI-MS/MS

<sup>g</sup> spot partially in gel front

**Table 3 : Functional categories of identified proteins**

<b>functional category</b>	<b>47</b>
enzyme	13
function unknown	7
signal transduction	6
protein degradation	6
protein, mitochondrial	4
cytoskeleton	2
secreted protein	2
transformation	2
mRNA splicing	1
stress response	1
transcription	1
translation	1
transport	1

The 47 non-redundant proteins identified in the proteomics approach in T47D-r and T47D fall into different functional categories as listed.